NVIDIA Blackwell Ultra Fuels AI & HPC Innovation, Efficiency and Capability

NVIDIA's groundbreaking hardware technologies and AI are unlocking unprecedented computational power. At the NVIDIA GTC 2025, NVIDIA unveiled its Blackwell Ultra GPU designed for the "Age of Reasoning" at its 2025 GPU Technology Conference (GTC). AI accelerators like GPUs are well suited for AI training and inference due to parallel processing, which allows for many calculations to be performed simultaneously. Only 30% of the top 500 supercomputers relied on accelerated computing; today, 80% do. The Green 500 ranking of supercomputers by energy efficiency shows an even more pronounced trend.

AI reasoning models emulate how the brain thinks to render a conclusion, popularized by OpenAI's o1, Google's Gemini 2.0 Flash Thinking and DeepSeek's R1 A1 models. Reasoning models improve responses to queries and more powerful GPUs improve the performance of these models. Blackwell Ultra GPUs are the next generation of the evolution of the GB200 bolstered by more inference power horsepower, packing 50% FLOPS at 1.1 exaFLOPS of FP dense compute.

At the NVIDIA GTC 2025, in his March 18 presentation titled, "The Next Frontier of AI Supercomputing: Efficiency with Unprecedented Capability", NVIDIA's Vice President of Hyperscale and HPC Computing, Ian Buck, stated, "Blackwell Ultra takes GB200's 40x data center revenue opportunity to 50x", citing faster token serving and higher throughput ideal for post-training for models like DeepSeek, which chomp through 100 trillion tokens.

Read more about the NVIDIA stock price prediction 2030.